



Kunutsor, S. K., Beswick, A. D., Whitehouse, M. R., & Blom, A. W. (2018). Systematic review of risk prediction scores for venous thromboembolism following joint replacement. *Thrombosis Research*, 168, 148-155. <https://doi.org/10.1016/j.thromres.2018.06.024>

Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/j.thromres.2018.06.024](https://doi.org/10.1016/j.thromres.2018.06.024)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at <https://www.sciencedirect.com/science/article/pii/S0049384818304080> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Systematic review of risk prediction scores for venous thromboembolism following joint replacement

Setor K. Kunutsor^{a,b,*}, Andrew D. Beswick^b, Michael R. Whitehouse^{a,b}, Ashley W. Blom^{a,b}

^aNational Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and University of Bristol, Bristol, UK

^bTranslational Health Sciences, Bristol Medical School, Musculoskeletal Research Unit, University of Bristol, Learning & Research Building (Level 1), Southmead Hospital, Bristol, BS10 5NB, UK

***Corresponding author at:**

Setor K. Kunutsor, Translational Health Sciences, Bristol Medical School, Musculoskeletal Research Unit, University of Bristol, Learning & Research Building (Level 1), Southmead Hospital, Bristol, BS10 5NB, UK; Phone: +44-7539589186; Fax: +44-1174147924;

Email address: setor.kunutsor@bristol.ac.uk

Abstract

Background: Venous thromboembolism (VTE) is an important cause of morbidity and a preventable cause of deaths following lower limb joint replacement. Risk prediction scores that help to predict individual VTE risk following lower limb joint replacement may inform the development of preventive strategies and guide treatment decisions. We aimed to systematically review the evidence on the development and/or validation of risk prediction scores for VTE following lower limb joint replacement.

Methods: Population-based studies that have developed and/or validated a risk prediction score for VTE following hip or knee replacement and randomized controlled trials (RCTs) that have evaluated the clinical impact of a score were searched for in MEDLINE, Embase, Web of Science, and The Cochrane Library to April 2018.

Results: Five observational cohort studies describing five risk scores were included. No RCTs were identified. The number of component variables in a single risk score ranged from 5 to 26. Two risk scores comprised 5-8 component variables. None of the studies reported calibration or discrimination statistics. Two risk scores were externally validated in single-institution cohorts and were reported to perform well. One study evaluated the general surgery Caprini risk score in primary hip and knee replacement patients and did not find it useful for VTE risk stratification.

Conclusions: Few VTE risk prediction scores in lower limb joint replacement exist and these have methodological issues, have been inadequately reported, not been sufficiently validated, and their impact on patient outcomes and decision making is unknown. Research is urgently warranted in the field.

Systematic review registration: PROSPERO 2018: CRD42018088712

Keywords: risk score; risk prediction; venous thromboembolism; deep vein thrombosis; pulmonary embolism; joint replacement; systematic review

1. Introduction

Total joint replacement, one of the most common elective orthopaedic procedures, is a highly successful and cost-effective intervention for alleviating pain and disability associated with advanced joint disease such as osteoarthritis.[1, 2] Venous thromboembolism (VTE), which comprises pulmonary embolism (PE) and deep vein thrombosis (DVT), is a frequent complication of lower limb joint replacement. Venous thromboembolism affects several millions of people globally, it is an important cause of long-term morbidity and a preventable cause of deaths, and its management is associated with huge health costs.[3] Despite the effectiveness of anticoagulants at preventing VTE, rates ranging from 0.27% to 61.0% have been reported in patient populations undergoing lower limb replacement.[4, 5]

With increasing life expectancy, there is a predicted large rise in the number of people who will be affected by joint disease and hence the number of patients undergoing primary joint replacement.[6] Despite the emergence of newer and more potent prophylactic regimens for VTE, there will also be a proportionate rise in the number of patients who will be affected by VTE. As such, the most appropriate way to tackle this is from a public health perspective – using a preventive approach. This entails identifying patients who are at high risk of developing VTE before they undergo joint replacement and developing preventative measures targeted at these high-risk groups. There are several known predisposing risk factors for VTE development following lower limb joint replacement and these include advanced age; high body mass index (BMI); smoking; and comorbidities such as cardiovascular disease, previous VTE, and cancer.[7-9] These established risk factors have the potential to be used to identify patients who are at high risk of VTE and can also be combined within a risk prediction score or prognostic model to predict VTE outcome risk for individuals. A risk prediction score is a statistical equation that uses multiple prognostic or risk factors in a formal combination to estimate the individualised probability or risk that a certain condition or disease will occur in the future.[10] Given that the risk of VTE development after lower limb joint replacement varies between individuals, there is an interest in developing individualised risk prediction scores for

VTE risk; however, little progress has been made in the area. Much research has been focussed on the identification of risk or prognostic factors for VTE in joint replacement patients.[11] Published studies have mostly reported on measures of the strength of the association (e.g., odds ratios, risk ratios, hazard ratios), which do not address the accuracy of these factors in classifying or predicting risk of VTE in individuals following joint replacement.[12]

Prevention of VTE following lower limb replacement is a high policy priority and no single risk prediction score has as yet been recommended as being optimal for VTE risk prediction in orthopaedic practice. It is also known that there is often conflicting evidence about the predictive performance of developed risk prediction scores.[13] To our knowledge, there is no summarised evidence on existing risk scores (including their component variables), their predictive performance, and whether their clinical effectiveness have been assessed in well-designed randomized controlled trials (RCTs). In this context, we aimed to systematically review all the available evidence on risk prediction scores for VTE following hip and/or knee replacement. The specific objectives were to: (i) assess clinical variables selected for model inclusion and the predictive performance of these models; (ii) assess if identified models have been externally validated and their performances compared; (iii) assess if the clinical effectiveness of these scores have been evaluated in appropriate RCTs; and (iv) to identify gaps in the existing evidence and whether further research is needed in the field. Our findings should inform clinical practice by identifying host, surgical, and laboratory characteristics that show consistent evidence of prognostic significance and should inform further research in this area.

Methods

2.1. Data sources and search strategy

We conducted this review in accordance with the CHARMS checklist[14] and PRISMA guidelines,[15] (**Appendix A**) and using a predefined protocol, which has been registered in the PROSPERO prospective register of systematic reviews (CRD42018088712). We searched for eligible

studies in MEDLINE, Embase, Web of Science, and The Cochrane Library electronic databases from inception up to 20 April 2018. The computer-based search strategy combined free and MeSH search terms and combination of key words relating to risk prediction (e.g., “risk score”, “sensitivity”, “prognostic model”), VTE (e.g., “venous thromboembolism”, “deep vein thrombosis”, “pulmonary embolism”), and joint replacement (e.g., “hip replacement”, “knee replacement”, “joint arthroplasty”). There were no language restrictions. We complemented the search by manually scanning reference lists of relevant articles and reviews for all relevant additional studies missed by the computerised search strategy. Details of the search strategy are provided in **Appendix B**.

2.2. Eligibility criteria

The following inclusion criteria were employed for eligibility of studies: (i) population-based observational studies (prospective cohort, retrospective cohort, prospective or retrospective case control, case-cohort, or nested-case control) that developed or validated a risk prediction score for VTE or an update on a previously developed model; (ii) outcome was VTE (DVT or PE) reported in a longitudinal design (Since 90% of VTE cases occur within the first post-operative week,[16] we considered a minimum of 30 days follow-up duration as acceptable); (iii) included adults > 18 years who had been followed up after hip and/or knee replacement surgery; and (iv) RCTs that assessed the clinical effectiveness of a VTE risk prediction score in an intervention group compared to usual care in a control group. We excluded the following: (i) cross-sectional and clinical case studies; (ii) studies with risk prediction scores containing less than two variables; (iii) studies only reporting measures of associations between a potential risk factor and the risk of VTE; and (iv) studies conducted in non-population-based samples.

2.3. Data extraction and quality assessment

One reviewer (S.K.K.) did an initial screen of titles and abstracts and acquired potentially relevant articles for full text evaluation. In any instance where there was confusion regarding eligibility of an

article based on title and abstract, that article was always selected for full text evaluation. Two independent reviewers (S.K.K., A.D.B.) assessed each article using the inclusion criteria and any disagreements regarding eligibility of an article was discussed, and consensus reached with involvement of a third reviewer (M.R.W). Data from selected articles was extracted by one reviewer (S.K.K.) using a standardized data collection form and quality assessments were also conducted. A second reviewer (A.D.B.) independently checked these data with that in original articles. We extracted data on first author's name, study publication date, country and geographical location, study design, baseline year, type of population, statistical model or methods employed, sample size, VTE outcomes, use of thromboprophylaxis, timing of outcomes, component variables of each risk prediction model, measures of discrimination, calibration, and/or reclassification, and reported performance comparisons of the model. We also extracted data on details of validation (internal and/or external) performance statistics. The risk of bias (quality) of any study developing or evaluating a risk prediction score was assessed using a preliminary version of the Prediction study Risk Of Bias Assessment Tool (PROBAST), a tool for assessing risk of bias, applicability and usability of prognostic models.[17] Briefly, this tool uses information on five pre-defined domains namely: participant selection, predictors, outcomes, sample size and patient flow, and analysis. The PROBAST evaluation is used to determine the risk of bias of the risk score (i.e., whether the score is likely to work as intended for the population of interest), with risk scores classified as low, moderate or high risk of bias.

2.4. Data synthesis and analysis

It was planned to quantitatively summarise the predictive performance (using calibration and discriminatory statistics) of models across studies if multiple studies were found to have validated the same risk score. This was to summarise the models' average performances across different settings and potential performance in a future setting. However, this was not possible given the limited number of studies, type of measures reported, unavailability of relevant summary statistics, and the

diversity of the study designs and populations. We were also unable to make effective comparisons of risk scores across studies because of the heterogeneity of the data and the variable methodologies adopted. The characteristics of each study and risk scores were summarized in tables. A narrative synthesis was performed using the data extracted and according to previously reported quality criteria for risk scores such as usability (10 or fewer components), good calibration, discriminative ability (> 0.70), generalizability (externally validated), and clinical effectiveness.[18, 19]

3. Results

3.1. Study identification and selection

The flow of studies through the screening and selection process is shown in **Figure 1**. The literature search strategy identified 603 potentially relevant articles. After an initial screen of titles and abstracts, 11 articles were selected for full text evaluation. Following detailed evaluation, six articles were excluded because (i) they did not report development and/or validation of a specific risk prediction score ($n=4$); (ii) the population was not relevant ($n=1$); and (iii) the outcome was not relevant ($n=1$). The remaining five articles met the inclusion criteria and were included in the review.[20-24]

3.2. Study characteristics and quality assessment

Table 1 summarizes characteristics of studies included in the review. Identified studies were published between 2012 and 2017. The sample size of cohorts ranged from 272 to 1,721,806 hip or knee replacements. Overall, the studies comprised 1,867,318 primary or revision hip and knee replacements, including 16,675 VTE events. Four studies were conducted in the USA and one in France. The average baseline age of participants ranged from 66.5 to 71.6 years for the three studies that reported these data. All included studies employed observational cohort designs and sampling frames included medical charts, institutional databases, national databases, and joint registries. Outcomes reported were VTE in 2 studies, PE in two studies and DVT in one study and all VTE cases

were symptomatic. Two studies reported on how the outcome event was defined, diagnosed and ascertained.[20, 22] Three studies reported use of post-operative pharmacological VTE thromboprophylaxis in patients which comprised a variety of agents including low-molecular-weight heparin (LMWH), warfarin, aspirin, and direct thrombin or factor Xa inhibitors (e.g., dabigatran, rivaroxaban, apixaban, etc).[20, 23, 24] Two studies reported having no records of the type of VTE prophylaxis received;[21, 22] however, one of these studies validated the risk score externally in a different population of patients who received warfarin[22] (**Table 2**). Quality assessment using the PROBAST tool showed evidence of high overall risk of bias throughout the included studies. Three risk scores had unclear concern for overall applicability and none of the scores were assessed as usable in the targeted individuals and context (**Appendix C**).

3.3. Model description and development

Table 2 provides details of VTE risk prediction scores described in included reports, statistical properties, their component variables, any measures of performance recorded, and reports of any validation and performance comparisons made. A total of five risk scores were described in the five eligible studies. Three risk scores were derivations of risk models on a base population[20-22] and two studies attempted to validate existing risk scores that had been developed in different populations or used different outcomes.[23, 24] None of the risk models was developed in a cohort designed for this sole purpose; they all retrospectively used data that had been prospectively collected for different purposes. Two studies employed logistic regression techniques in their model development[20, 21] and one employed Cox regression.[22] The component risk factors varied from score to score and ranged in number from 5 to 26. The component variables for the risk scores were based on data that could be assessed non-invasively such as demographics, BMI, medical and surgical histories, and surgical procedures.

3.4. Model diagnostics

None of studies reported measures of discrimination or calibration for any of the risk scores described. Parvizi and colleagues created a nomogram based on eight risk factors to predict the risk of PE after hip and knee replacement.[20] However, the study authors did not report on the performance of the nomogram. In another study, Parvizi and colleagues used National Inpatient Sample (NIS) registry data to develop an individualised risk model for VTE which was based on 26 risk factors.[21] The authors used scoring criteria to assess the performance of the model. The authors reported that their calibration curve showed a near perfect fit between the predicted VTE rate (using the risk model) and the actual rate of VTE in NIS data up to a 5% rate of VTE, beyond which point there was a clear divergence. Bohl and colleagues developed a risk stratification system for PE within 30 days of primary total hip or knee replacement, using the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) registry data.[22] The risk stratification score, which comprised of five simple patient variables, was based on a point-scoring system in which patients were assigned as low-, medium-, and high-risk categories. Bateman and colleagues[23] evaluated the Caprini Risk Assessment Model (RAM), which was originally developed to estimate VTE risk in the general surgical population.[25] The authors reported that this score did not provide clinically useful risk stratification for total hip and knee replacement patients. Dauty and colleagues used the Risk Assessment and Predictor Tool (RAPT) to evaluate the risk of complications (including DVT) in patients following total knee replacement surgery.[24] The RAPT tool is a validated risk score which was originally developed to predict a patient's risk of needing extended inpatient rehabilitation after hip or knee replacement.[26] Findings from the study suggested that the RAPT tool was appropriate in identifying patients who had the most complications (including DVT) and required a longer hospital stay.

3.5. Model validation

Only one study reported conducting internal validation using bootstrapping.[20] Two risk scores were externally validated using independent datasets in the same studies. The risk score based on NIS data was reported to perform well when compared with the independent dataset.[21] The ACS-NSQIP-derived risk stratification system was externally validated in a single-institution cohort and was observed to perform well[22] (**Table 2**).

3.6. Performance comparisons

None of the risk scores were compared with existing models.

3.7. Clinical evaluation of risk scores

None of the studies described the evaluation of the clinical effectiveness of a risk prediction score in an intervention study or as part of an impact study aimed at changing patient outcomes.

4. Discussion

4.1. Key findings

This systematic review of available risk prediction scores for VTE following hip and knee replacement identified five studies, all published in the last six years. Three studies reported the development of three independent risk scores,[20-22] whereas two studies evaluated existing scores originally developed for different populations or outcomes.[23, 24] Only two studies reported on the definitions used for VTE outcomes.[20, 22] The number of component variables for the five risk prediction scores ranged from 5 to 26, with three scores having less than 10 components. None of the studies reported appropriate calibration or discrimination statistics for the risk scores; however, one study reported on a calibration curve which suggested good performance of the model below a specific VTE rate.[21] Two of the risk scores were externally validated in independent populations by

the authors in the same studies.[21, 22] Quality assessment of the risk scores' development and validation criteria showed all scores to have a high overall risk of bias. This was mainly due to the methodology used in assessment of predictors and outcomes, as well as the analyses employed.

4.2. Methodological limitations

It is evident that there has been little progress in the field of development of risk prediction scores for VTE following hip or knee replacement. In addition to the huge gaps in the field, our findings highlight the methodological limitations in the development of the identified scores. First, the majority of studies did not report on the definitions used for VTE outcomes. This tends to create uncertainty as to which populations the proposed risk scores are applicable to. In addition, VTE outcomes that are not independently adjudicated and based on database International Classification of Diseases (ICD) codes that are not validated, are less accurate. To ensure that risk prediction scores are reliable for the intended patients, standard consistent definitions need to be employed.[27] Second, of the three studies that reported development of a risk prediction model, none was developed in a cohort recruited for this sole purpose, thereby introducing an inherent selection bias. The study designs employed retrospective data collected from prospective cohort designs, which are ideal for risk score modelling as the risk factor information can be ascertained blindly in relation to the outcome or disease.[27] Third, a key methodological issue was that none of the studies reported calibration, discrimination, as well as reclassification statistics, which describe the performance of the models. This is of great concern given that these measures are very vital in the development of a risk score. Calibration, which is measured by the goodness-of-fit statistic, is the ability to accurately estimate the risk of a future event (VTE in this case). Discrimination is the ability of the risk prediction score to separate individuals at higher risk from those at lower risk of the event of interest; this is assessed using the area under the receiver-operating characteristic (ROC) curve or C statistic (Harrell's C-index[28]).[29] A C-index of 0.5 represents no improvement over what would be expected by chance, while a C-index of 1 implies perfect discrimination.[30] Reclassification or risk stratification, which

is assessed using the net-reclassification-improvement (NRI)[31, 32] and the integrated-discrimination-improvement (IDI),[31] is the ability of the risk score to appropriately reclassify patients into clinically relevant subgroups. These two measures test the addition of factors to risk prediction models. How well these identified risk scores are calibrated and their discriminative abilities are not known given the lack of relevant information provided in the studies. Fourth, none of the studies provided a clear and detailed report on the treatment of missing data, an issue of relevance in the development of risk scores.[10] Bohl and colleagues reported excluding patients with missing data on relevant variables.[22] Evidence suggests that risk prediction scores that employ multiple imputation technique produce more generalisable models compared with models that ignore such additional analyses.[27] Fifth, given that simplicity is an important criteria for developing clinically useful risk prediction scores,[33, 34] there are concerns with usability of two of the risk scores, as they employed 20-26 variables.[21, 23] Complex models are more likely to provide overoptimistic predictions, especially when extensive variable selection has been performed.[35] Sixth, except for one study,[20] none of the risk scores was reported to have undergone internal validation, which is a process which provides a good indication of how optimistic the risk score may be.[36] Nevertheless, despite the importance of internal validation, this process is unable to provide information on the performance of the risk score elsewhere or its generalizability. Before a risk prediction tool can be effectively employed in clinical practice or in a real world setting, its generalization needs evaluation in a new setting (using independent datasets from different locations) – a process known as external validation.[10] Two of the studies externally validated their developed risk scores;[21, 22] however, none of these risk scores were reported to have been externally validated by a third party, a method which is least prone to bias.[37] while a third study attempted to validate an existing risk score which was not specifically developed for that patient population. Finally, none of the identified risk scores was reported to have been evaluated in an impact study to investigate their influence on patient outcomes, decision making, and costs, a vital criterion that also needs to be fulfilled before a risk score can be implemented in a clinical setting.[10] For a risk score to be adopted in clinical practice, it

should be clinically credible, well calibrated with good discriminative ability, should have been externally validated, and have real clinical impact.[10, 33, 38]

4.3. Implications of our findings

Findings from our review suggest the potential value of VTE risk assessment is underappreciated in orthopaedic practice. Individualized patient risk assessment to identify VTE risk in surgical patients has been very widely advocated.[39-41] Indeed, several risk assessment models have been developed to predict VTE risk in surgical patients, but it appears none have been extensively validated or clinically evaluated in hip or knee replacement patients. The most widely used VTE risk assessment tool in surgical practice is the Caprini Risk Assessment Model (RAM), first described almost three decades ago.[25] The original Caprini RAM integrates 20 risk factors (comprising age, genetics, medical, and surgical factors) to calculate a cumulative risk score, which is used to group patients into risk categories (low, moderate, and high risk). It has since then undergone several modifications and refinements as well as validation in both medical and surgical patients.[42-47] In one of the studies included in our review, the study authors employed the Caprini RAM in a first attempt to validate it in total hip and knee replacement patients. However, the risk model was unhelpful in stratifying VTE risk, which the authors attributed to inadequate power to validate. Among the identified risk scores, the ACS-NSQIP-derived risk stratification system may be potentially promising for future clinical use based on its development from a large dataset, use of five patient characteristics, and having undergone external validation. However, in addition to the methodological limitations and lack of information on its discriminative ability, its clinical effectiveness is yet to be evaluated and needs to undergo further validation in new populations. The limited number of published VTE risk scores available for hip or knee replacement patients and lack of data on their performance is a cause for great concern. The incidence of VTE is likely to increase in conjunction with growing healthcare burden due to osteoarthritis[48] and a predicted large rise in the numbers of hip and knee replacement procedures.[49, 50] In the era of preventive medicine, VTE risk assessment should be performed in all

patients undergoing lower limb replacement using reliable and well validated risk prediction scores. Our findings suggest validated tools are currently non-existent. Researchers and clinicians are encouraged to support collaborative efforts to develop and validate appropriate risk prediction scores for use in orthopaedic practice.

4.3. Study strengths and limitations

To our knowledge, this is the first systematic review identifying and summarizing the literature on the development and validation of VTE risk prediction scores following hip or knee replacement. Though we retrieved a limited number of eligible studies, we employed a comprehensive search strategy which spanned multiple databases, making it unlikely that any relevant study was missed. Indeed, this study has revealed very large gaps in the research area. We also conducted a detailed assessment of the existing risk scores using established criteria. A limitation of our review was that we were unable to harmonize data from contributing studies to perform pooled analysis and analysis by subgroups such as age, sex, type of joint (hip vs knee), type of VTE (DVT vs PE, symptomatic vs asymptomatic) and type of anticoagulant; this was due to the limited number of studies, lack of data on measures of performance (e.g., C-index), and substantial heterogeneity between studies. The heterogeneity between the included studies was attributable to the different VTE outcomes reported, different thromboprophylactic agents used, different percentage of patients with outcomes (ranging from <1% to 6%) and different follow-up durations. Despite efforts to summarise the data as robustly as possible using established criteria, our conclusions might be limited due to the quality of published research and the inability of studies to report the results in a way that can be used by clinicians.

5. Conclusions

Only a small number of risk scores to predict VTE in hip and knee joint replacement patients have been developed and these have several limitations. The existing risk scores have been developed using inadequate methodology, have been inadequately reported, not been sufficiently validated, and their

impact on patient outcomes and decision making is unknown. The ACS-NSQIP-derived risk stratification system may have some potential for use in clinical practice; however, inadequate methodology was employed in its development, there is lack of detailed information on its performance, it has not undergone further validation in new populations, and its clinical impact hasn't been evaluated. Urgent research is encouraged in the field to help develop robust risk prediction scores with potential clinical value.

Funding sources

This study was supported by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care.

Declarations of interest

None.

References

- [1] P.R. Fortin, A.E. Clarke, L. Joseph, M.H. Liang, M. Tanzer, D. Ferland, C. Phillips, A.J. Partridge, P. Belisle, A.H. Fossel, N. Mahomed, C.B. Sledge, J.N. Katz, Outcomes of total hip and knee replacement: preoperative functional status predicts outcomes at six months after surgery, *Arthritis and rheumatism* 42(8) (1999) 1722-8.
- [2] M.A. Ritter, M.J. Albohm, E.M. Keating, P.M. Faris, J.B. Meding, Comparative outcomes of total joint arthroplasty, *J Arthroplasty* 10(6) (1995) 737-41.
- [3] L. Li, P. Zhang, J.H. Tian, K. Yang, Statins for primary prevention of venous thromboembolism, *Cochrane Database Syst Rev* 12 (2014) CD008203.
- [4] B.F. Gage, A.R. Bass, H. Lin, S.C. Woller, S.M. Stevens, N. Al-Hammadi, J. Li, T. Rodriguez, Jr., J.P. Miller, G.A. McMillin, R.C. Pendleton, A.K. Jaffer, C.R. King, B.D. Whipple, R. Porche-Sorbet, L. Napoli, K. Merritt, A.M. Thompson, G. Hyun, J.L. Anderson, W. Hollomon, R.L. Barrack, R.M. Nunley, G. Moskowitz, V. Davila-Roman, C.S. Eby, Effect of Genotype-Guided Warfarin Dosing on Clinical Events and Anticoagulation Control Among Patients Undergoing Hip or Knee Arthroplasty: The GIFT Randomized Clinical Trial, *JAMA* 318(12) (2017) 1115-1124.
- [5] H. Zhang, P. Mao, C. Wang, D. Chen, Z. Xu, D. Shi, J. Dai, Y. Yao, Q. Jiang, Incidence and risk factors of deep vein thrombosis (DVT) after total hip or knee arthroplasty: a retrospective study with routinely applied venography, *Blood Coagul Fibrinolysis* 28(2) (2017) 126-133.
- [6] S.M. Kurtz, E. Lau, K. Ong, K. Zhao, M. Kelly, K.J. Bozic, Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030, *Clin Orthop* 467(10) (2009) 2606-12.
- [7] N.F. Soohoo, E. Farnig, J.R. Lieberman, L. Chambers, D.S. Zingmond, Factors that predict short-term complication rates after total hip arthroplasty, *Clin Orthop* 468(9) (2010) 2363-71.
- [8] S.G. Memtsoudis, M.C. Besculides, L. Gaber, S. Liu, A. Gonzalez Della Valle, Risk factors for pulmonary embolism after hip and knee arthroplasty: a population-based study, *Int Orthop* 33(6) (2009) 1739-45.
- [9] R.H. White, M.C. Henderson, Risk factors for venous thromboembolism after total hip and knee replacement surgery, *Curr Opin Pulm Med* 8(5) (2002) 365-71.
- [10] E.W. Steyerberg, K.G. Moons, D.A. van der Windt, J.A. Hayden, P. Perel, S. Schroter, R.D. Riley, H. Hemingway, D.G. Altman, P. Group, Prognosis Research Strategy (PROGRESS) 3: prognostic model research, *PLoS Med* 10(2) (2013) e1001381.
- [11] J. Zhang, Z. Chen, J. Zheng, S.J. Breusch, J. Tian, Risk factors for venous thromboembolism after total hip and total knee arthroplasty: a meta-analysis, *Arch Orthop Trauma Surg* 135(6) (2015) 759-72.
- [12] M.J. Pencina, R.B. D'Agostino, R.S. Vasan, Statistical methods for assessment of added usefulness of new biomarkers, *Clin Chem Lab Med* 48(12) (2010) 1703-11.
- [13] T.P. Debray, J.A. Damen, K.I. Snell, J. Ensor, L. Hooft, J.B. Reitsma, R.D. Riley, K.G. Moons, A guide to systematic review and meta-analysis of prediction model performance, *BMJ* 356 (2017) i6460.

- [14] K.G. Moons, J.A. de Groot, W. Bouwmeester, Y. Vergouwe, S. Mallett, D.G. Altman, J.B. Reitsma, G.S. Collins, Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist, *PLoS Med* 11(10) (2014) e1001744.
- [15] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, *PLoS Med* 6(7) (2009) e1000097.
- [16] J. Parvizi, R. Huang, I.J. Raphael, M.G. Maltenfort, W.V. Arnold, R.H. Rothman, Timing of Symptomatic Pulmonary Embolism with Warfarin Following Arthroplasty, *J Arthroplasty* 30(6) (2015) 1050-3.
- [17] Prediction model study Risk Of Bias Assessment Tool (PROBAST). Available from: <http://www.systematic-reviews.com/probast/>. Accessed on 12 July, 2016.
- [18] D.G. Altman, Y. Vergouwe, P. Royston, K.G. Moons, Prognosis and prognostic research: validating a prognostic model, *BMJ* 338 (2009) b605.
- [19] D. Noble, R. Mathur, T. Dent, C. Meads, T. Greenhalgh, Risk models and scores for type 2 diabetes: systematic review, *BMJ* 343 (2011) d7163.
- [20] J. Parvizi, R. Huang, I.J. Raphael, W.V. Arnold, R.H. Rothman, Symptomatic pulmonary embolus after joint arthroplasty: stratification of risk factors, *Clin Orthop* 472(3) (2014) 903-12.
- [21] J. Parvizi, R. Huang, M. Rezapoor, B. Bagheri, M.G. Maltenfort, Individualized Risk Model for Venous Thromboembolism After Total Joint Arthroplasty, *J Arthroplasty* 31(9 Suppl) (2016) 180-6.
- [22] D.D. Bohl, M.G. Maltenfort, R. Huang, J. Parvizi, J.R. Lieberman, C.J. Della Valle, Development and Validation of a Risk Stratification System for Pulmonary Embolism After Elective Primary Total Joint Arthroplasty, *J Arthroplasty* 31(9 Suppl) (2016) 187-91.
- [23] D.K. Bateman, R.W. Dow, A. Brzezinski, H.Y. Bar-Eli, S.T. Kayiaros, Correlation of the Caprini Score and Venous Thromboembolism Incidence Following Primary Total Joint Arthroplasty-Results of a Single-Institution Protocol, *J Arthroplasty* 32(12) (2017) 3735-3741.
- [24] M. Dauty, X. Schmitt, P. Menu, B. Rousseau, C. Dubois, Using the Risk Assessment and Predictor Tool (RAPT) for patients after total knee replacement surgery, *Ann Phys Rehabil Med* 55(1) (2012) 4-15.
- [25] J.A. Caprini, J.I. Arcelus, J.H. Hasty, A.C. Tamhane, F. Fabrega, Clinical assessment of venous thromboembolic risk in surgical patients, *Semin Thromb Hemost* 17 Suppl 3 (1991) 304-12.
- [26] L.B. Oldmeadow, H. McBurney, V.J. Robertson, Predicting risk of extended inpatient rehabilitation after hip or knee arthroplasty, *J Arthroplasty* 18(6) (2003) 775-9.
- [27] J. Ensor, R.D. Riley, D. Moore, K.I. Snell, S. Bayliss, D. Fitzmaurice, Systematic review of prognostic models for recurrent venous thromboembolism (VTE) post-treatment of first unprovoked VTE, *BMJ open* 6(5) (2016) e011190.
- [28] F.E. Harrell, Jr., K.L. Lee, D.B. Mark, Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, *Stat Med* 15(4) (1996) 361-87.

- [29] N.R. Cook, Use and misuse of the receiver operating characteristic curve in risk prediction, *Circulation* 115(7) (2007) 928-35.
- [30] F.E. Harrell, Jr., R.M. Califf, D.B. Pryor, K.L. Lee, R.A. Rosati, Evaluating the yield of medical tests, *JAMA* 247(18) (1982) 2543-6.
- [31] M.J. Pencina, R.B. D'Agostino, Sr., R.B. D'Agostino, Jr., R.S. Vasan, Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond, *Stat Med* 27(2) (2008) 157-72; discussion 207-12.
- [32] M.J. Pencina, R.B. D'Agostino, Sr., E.W. Steyerberg, Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers, *Statistics in Medicine* 30(1) (2011) 11-21.
- [33] J.C. Wyatt, D.G. Altman, Commentary: Prognostic models: clinically useful or quickly forgotten?, *BMJ* 311(7019) (1995) 1539-1541.
- [34] K.G. Moons, P. Royston, Y. Vergouwe, D.E. Grobbee, D.G. Altman, Prognosis and prognostic research: what, why, and how?, *BMJ* 338 (2009) b375.
- [35] W. Sauerbrei, The Use of Resampling Methods to Simplify Regression Models in Medical Statistics, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 48(3) (1999) 313-329.
- [36] E.W. Steyerberg, S.E. Bleeker, H.A. Moll, D.E. Grobbee, K.G. Moons, Internal and external validation of predictive models: a simulation study of bias and precision in small samples, *J Clin Epidemiol* 56(5) (2003) 441-7.
- [37] S.R. Barber, M.J. Davies, K. Khunti, L.J. Gray, Risk assessment tools for detecting those with pre-diabetes: a systematic review, *Diabetes Res Clin Pract* 105(1) (2014) 1-13.
- [38] B.M. Reilly, A.T. Evans, Translating clinical research into clinical practice: impact of using prediction rules to make decisions, *Ann Intern Med* 144(3) (2006) 201-9.
- [39] G. Agnelli, G. Bolis, L. Capussotti, R.M. Scarpa, F. Tonelli, E. Bonizzoni, M. Moia, F. Parazzini, R. Rossi, F. Sonaglia, B. Valarani, C. Bianchini, G. Gussoni, A clinical outcome-based prospective study on venous thromboembolism after cancer surgery: the @RISTOS project, *Ann Surg* 243(1) (2006) 89-95.
- [40] D.A. Hatef, J.M. Kenkel, M.Q. Nguyen, J.P. Farkas, F. Abtahi, R.J. Rohrich, S.A. Brown, Thromboembolic risk assessment and the efficacy of enoxaparin prophylaxis in excisional body contouring surgery, *Plast Reconstr Surg* 122(1) (2008) 269-79.
- [41] NICE Clinical guideline NG89: Venous thromboembolism in over 16s: reducing the risk of hospital-acquired deep vein thrombosis or pulmonary embolism. March 2018. Available at: <https://www.nice.org.uk/guidance/ng89/chapter/Recommendations#risk-assessment>. Accessed on 25 April 2018.
- [42] J.A. Caprini, Thrombosis risk assessment as a guide to quality patient care, *Dis Mon* 51(2-3) (2005) 70-8.
- [43] J.A. Caprini, J.I. Arcelus, J.J. Reyna, Effective risk stratification of surgical and nonsurgical patients for venous thromboembolic disease, *Semin Hematol* 38(2 Suppl 5) (2001) 12-9.

- [44] C.J. Pannucci, S.H. Bailey, G. Dreszer, C. Fisher Wachtman, J.W. Zumsteg, R.M. Jaber, J.B. Hamill, K.M. Hume, J.P. Rubin, P.C. Neligan, L.K. Kallianen, R.E. Hoxworth, A.L. Pusic, E.G. Wilkins, Validation of the Caprini risk assessment model in plastic and reconstructive surgery patients, *J Am Coll Surg* 212(1) (2011) 105-12.
- [45] V. Bahl, H.M. Hu, P.K. Henke, T.W. Wakefield, D.A. Campbell, Jr., J.A. Caprini, A validation study of a retrospective venous thromboembolism risk scoring method, *Ann Surg* 251(2) (2010) 344-50.
- [46] J.I. Arcelus, S. Candocia, C.I. Traverso, F. Fabrega, J.A. Caprini, J.H. Hasty, Venous thromboembolism prophylaxis and risk assessment in medical patients, *Semin Thromb Hemost* 17 Suppl 3 (1991) 313-8.
- [47] N.A. Zakai, J. Wright, M. Cushman, Risk factors for venous thrombosis in medical inpatients: validation of a thrombosis risk score, *J Thromb Haemost* 2(12) (2004) 2156-61.
- [48] M. Cross, E. Smith, D. Hoy, S. Nolte, I. Ackerman, M. Fransen, L. Bridgett, S. Williams, F. Guillemin, C.L. Hill, L.L. Laslett, G. Jones, F. Cicuttini, R. Osborne, T. Vos, R. Buchbinder, A. Woolf, L. March, The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study, *Ann Rheum Dis* 73(7) (2014) 1323-30.
- [49] S. Kurtz, K. Ong, E. Lau, F. Mowat, M. Halpern, Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030, *J Bone Joint Surg Am* 89-A(4) (2007) 780-785.
- [50] A. Patel, G. Pavlou, R.E. Mujica-Mota, A.D. Toms, The epidemiology of revision total knee and hip arthroplasty in England and Wales: a comparative analysis with projections for the United States. A study using the National Joint Registry dataset, *Bone Joint J* 97-B(8) (2015) 1076-81.

Figure legends

Figure 1. PRISMA Flow Diagram.

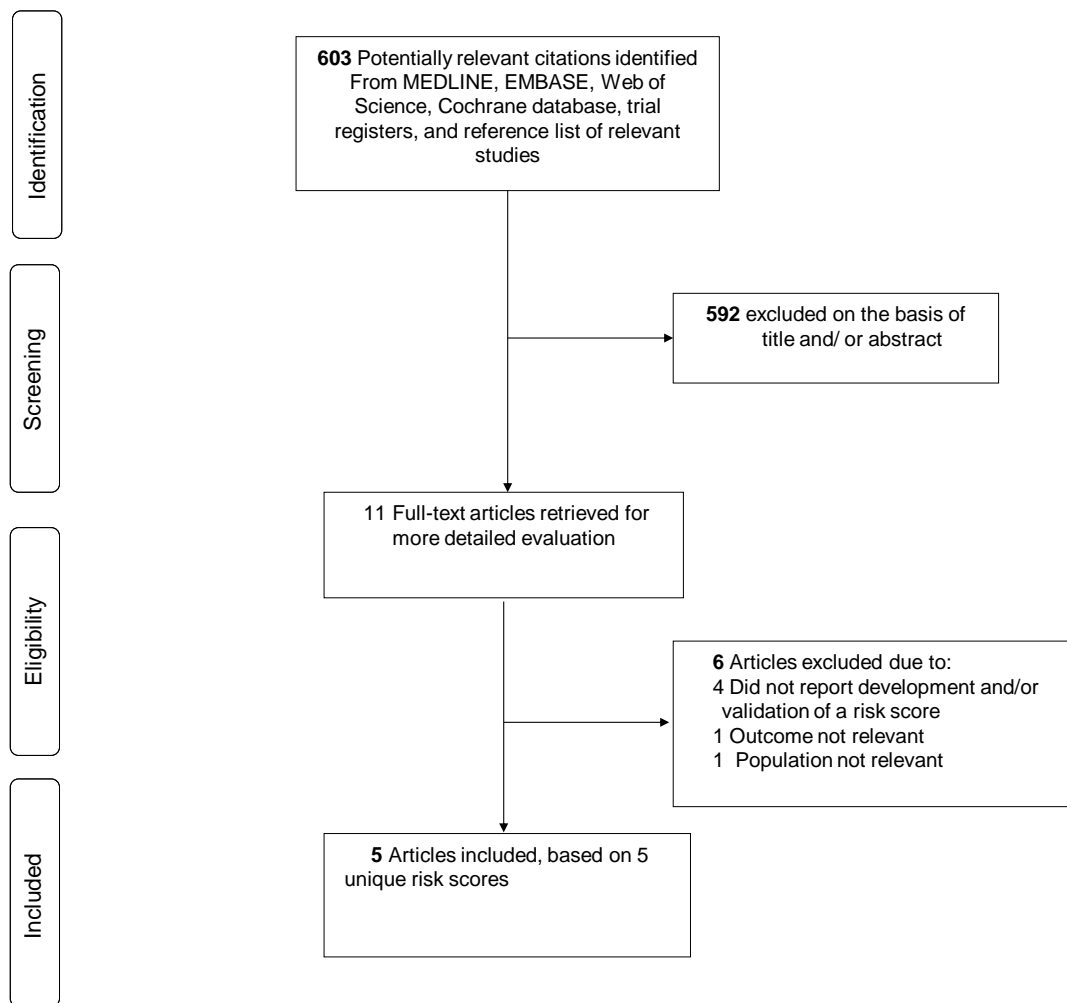


Table 1. Summary characteristics of studies included in the review

Lead author, publication date (reference)	Location	Baseline year	Study design	Sampling frame	Population	Anticoagulation and dosage	Mean age (years)	Name of risk tool	Specific outcome reported	Sample size (joint replacements or patients)	Number of events (%)	Timing of outcome
Dauty, 2012 (24)	France	2004-2007	Retrospective cohort	Medical charts	Total knee replacement	21-day LMWH in preventive enoxaparin sodium injections	71.6	RAPT	Symptomatic DVT	272	17 (6.25)	NR
Parvizi, 2014 (20)	USA	2000-2011	Observational cohort	Institutional orthopaedic database	Primary and revision total hip and knee replacement	Warfarin for 6 weeks or 325 mg aspirin twice daily	69.4	NR	Symptomatic PE	26,391	281 (1.06)	90 days
Parvizi, 2016 (21)	USA	2002-2011	Observational cohort	NIS data	Total hip and knee replacement	NR	40-100‡	NR	Symptomatic VTE (DVT and PE)	1,721,806	15,775 (0.92)	NR
Bohl, 2016 (22)	USA	2006-2013	Observational cohort	ACS-NSQIP	Primary total hip or knee replacement	NR	NR	ACS-NSQIP-derived risk stratification system	Symptomatic PE	118,473	~592† (0.50)	30 days
Bateman, 2017 (23)	USA	2015-2016	Retrospective cohort	Orthopaedic database	Primary total hip or knee replacement	Warfarin, direct thrombin or factor Xa inhibitors (e.g., dabigatran, apixaban, 10 mg rivaroxaban daily), or aspirin 325 mg twice daily	66.5	Caprini RAM*	Symptomatic VTE (DVT and PE)	376	10 (2.66)	90 days

ACS-NSQIP, American College of Surgeons National Surgical Quality Improvement Program; DVT, deep vein thrombosis; LMWH, low-molecular-weight heparin; NIS, National Inpatient Sample; NR, not reported; PE, pulmonary embolism; RAM, risk assessment model; RAPT, Risk Assessment and Predictor Tool; THA, VTE, venous thromboembolism

*, study employed the Caprini RAM in predicting 90-day postoperative VTE incidence in total joint replacement patients; †, numbers calculated based on percentage provided in the report; ‡, age range

Table 2. Key characteristics of development and validation of risk prediction scores for venous thromboembolism included in the review

Lead author, publication date	Name of risk tool	Statistical model	Predictors used	Number of predictors	Discrimination (C-index)	Calibration (HL goodness-of-fit test)	Internal validation	External validation	Performance comparison
Dauty, 2012 (24)	RAPT	NR	Age, gender, average walking distance, use of gait aid, use of community support and care, and social support at discharge	6	NR	NR	None	None	NR
Parvizi, 2014 (20)	NR	Logistic regression	Knee surgery, CCI, atrial fibrillation, postoperative DVT, COPD, anaemia, depression, BMI	8	NR	NR	Bootstrapping	None	NR
Parvizi, 2016 (21)	NR	Logistic regression	Bilateral joints, not primary THA, age, anaemia, CHF, lymphoma, fluid and electrolyte disorders, metastatic cancer, peripheral vascular disease, non-metastatic solid tumours, weight loss, chronic pulmonary heart disease, blood transfusion, history of VTE, myeloproliferative disorders, hypercoagulability state, myocardial infarction, varicose veins, fracture, inflammatory bowel disease, sepsis, periprosthetic joint infection, atrial fibrillation, stroke, apnoea	26	NR	Near perfect goodness of fit reported in NIS data up to 5% rate of VTE	None	Externally validated in a single-institution cohort	NR
Bohl, 2016 (22)	ACS-NSQIP-derived risk stratification system	Cox regression	Age, sex, BMI, preoperative haematocrit, and procedure type	5	NR	NR	None	Externally validated in a single-institution cohort in which all patients received warfarin as thromboprophylaxis	NR
Bateman, 2017 (23)	NR	NA	Age, planned operation >2 hours, history of DVT or PE, leg oedema/ulcers/ stasis, sepsis, varicose veins, hormone treatment, malignancy, previous immobilisation, CVD, trauma, fracture, obesity, stroke, major surgery, pregnancy, protein C/ antithrombin III/ protein S deficiency, plasminogen disorders, nephrotic syndrome, paroxysmal nocturnal haemoglobinuria, lupus, polycythaemia vera, inflammatory bowel disease, and other	20	NR	NR	None	Validation of the Caprini RAM in joint replacement patients	NR

ACS-NSQIP, American College of Surgeons National Surgical Quality Improvement Program; BMI, body mass index; CCI, Charlson Comorbidity Index; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; DVT, deep vein thrombosis; NA, not applicable; NR, not reported; RAPT, Risk Assessment and Predictor Tool; THA, total hip arthroplasty; VTE, venous thromboembolism

Supplementary Material

Appendix A	PRISMA checklist
Appendix B	Literature search strategy
Appendix C	Quality assessment of included risk scores using the PROBAST tool

Appendix A. PRISMA checklist

Section/topic	Item No	Checklist item	Reported on page No
Title			
Title	1	Identify the report as a systematic review, meta-analysis, or both	1
Abstract			
Structured summary	2	Provide a structured summary including, as applicable, background, objectives, data sources, study eligibility criteria, participants, interventions, study appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, systematic review registration number	2
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	3-4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)	5
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (such as web address), and, if available, provide registration information including registration number	5
Eligibility criteria	6	Specify study characteristics (such as PICOS, length of follow-up) and report characteristics (such as years considered, language, publication status) used as criteria for eligibility, giving rationale	5
Information sources	7	Describe all information sources (such as databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	Appendix 2
Study selection	9	State the process for selecting studies (that is, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	5
Data collection process	10	Describe method of data extraction from reports (such as piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	6
Data items	11	List and define all variables for which data were sought (such as PICOS, funding sources) and any assumptions and simplifications made	6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	6
Summary measures	13	State the principal summary measures (such as risk ratio, difference in means).	Not applicable
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (such as I^2 statistic) for each meta-analysis	Not applicable
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (such as publication bias, selective reporting within studies)	6
Additional analyses	16	Describe methods of additional analyses (such as sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified	Not applicable
Results			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram	7 and Fig. 1
Study characteristics	18	For each study, present characteristics for which data were extracted (such as study size, PICOS, follow-up period) and provide the citations	7-8, Table1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see item 12).	7-8, Appendix 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present for each study (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot	Not applicable
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency	Not applicable
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15)	Not applicable
Additional analysis	23	Give results of additional analyses, if done (such as sensitivity or subgroup analyses, meta-regression) (see item 16)	Not applicable
Discussion			
Summary of evidence	24	Summarise the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (such as health care providers, users, and policy makers)	10
Limitations	25	Discuss limitations at study and outcome level (such as risk of bias), and at review level (such as incomplete retrieval of identified research, reporting bias)	13
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research	10-14
Funding			
Funding	27	Describe sources of funding for the systematic review and other support (such as supply of data) and role of funders for the systematic review	14

Appendix B. Literature search strategy

Relevant studies, published before 20 April 2018 (date last searched), were identified through electronic searches not limited to the English language using MEDLINE, EMBASE, Web of Science, and Cochrane databases. Electronic searches were supplemented by scanning reference lists of articles identified for all relevant studies (including review articles) and by hand searching of relevant journals. The computer-based searches combined search terms related to risk prediction, venous thromboembolism, and joint replacement.

- 1 exp Venous Thrombosis/ (51140)
- 2 venous thrombus.mp. (493)
- 3 exp Venous Thromboembolism/ (8174)
- 4 exp Pulmonary Embolism/ (36029)
- 5 deep vein thrombosis.mp. (14494)
- 6 risk score.mp. (11236)
- 7 exp Risk Assessment/ (228856)
- 8 predict.mp. (286230)
- 9 score.mp. (449810)
- 10 diagnostic.mp. (737169)
- 11 exp "Sensitivity and Specificity"/ (522324)
- 12 exp ROC CURVE/ (46410)
- 13 receiver operating characteristic.mp. (43083)
- 14 exp Area Under Curve/ (35130)
- 15 C statistic.mp. (3339)
- 16 C-index.mp. (1879)
- 17 concordance statistic.mp. (105)
- 18 prognostic.mp. (238170)
- 19 exp Algorithms/ (273054)
- 20 model.mp. (1744458)
- 21 calculator.mp. (3012)
- 22 exp ARTHROPLASTY/ (56897)
- 23 exp Arthroplasty, Replacement/ (45495)
- 24 joint arthroplasty.mp. (3488)
- 25 total arthroplasty.mp. (359)
- 26 exp Arthroplasty, Replacement, Knee/ (19139)
- 27 exp Arthroplasty, Replacement, Hip/ (22802)
- 28 exp Arthroplasty, Replacement, Ankle/ (553)
- 29 exp Arthroplasty, Replacement, Shoulder/ (380)
- 30 1 or 2 or 3 or 4 or 5 (90506)
- 31 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 (3826985)
- 32 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 (58398)
- 33 30 and 31 and 32 (462)
- 34 limit 33 to (humans and "all adult (19 plus years)") (310)

Each part was specifically translated for searching the other databases (EMBASE, Web of Science, and Cochrane databases)

Appendix C. Quality assessment of included risk scores using the PROBAST tool

<i>Author, year of publication</i>	<i>Risk tool</i>	<i>Type of model evaluation</i>	<i>Participant selection</i>	<i>Predictors</i>	<i>Outcome</i>	<i>Sample size and participant flow</i>	<i>Analysis</i>	<i>Overall risk of bias</i>	<i>Overall applicability (concern)</i>	<i>Usability of model</i>
Dauty, 2012	RAPT	Development	-	-	-	-	-	-	-	No
Parvizi, 2014	NR	Development	?	?	?	-	-	-	-	No
Parvizi, 2016	NR	Development	+	?	-	-	-	-	-	No
Parvizi, 2016	NR	External validation	?	?	-	-	-	-	-	No
Bohl, 2016	ACS-NSQIP	Development	+	?	?	?	-	-	-	No
Bohl, 2016	ACS-NSQIP	External validation	+	?	?	?	-	-	-	No
Bateman, 2017	NR	Development	?	?	-	-	-	-	-	No
+	Low risk of bias									
?	Unclear risk of bias									
-	High risk of bias									

ACS-NSQIP, American College of Surgeons National Surgical Quality Improvement Program; NR, not reported